Universiteit Leiden
Governance and Global Affairs

Resignation Calls, Reallocations and Individual Ministerial Terminations in Presidential Democracies

Bastián González-Bustamante
Leiden University
✉ b.a.gonzalez.bustamante@fgga.leidenuniv.nl

Political Economy and Public Policy Talk at the Institute of Public Administration
The Hague, Netherlands, September 30, 2024

Table of Contents

Introduction

Introduction

My research question is: **How do resignation calls affect cabinet reallocations and dismissals of ministers in Latin American presidential democracies?**

I use a combination of a presidential **protection policy** and a **dismissal rule**: The president protects ministers subjected to a first questioning but, because they are then contaminated, the probability of ministerial termination increases as from a second call for their resignation.

- Archival review of press reports in 12 LA countries (mid-1970s-2021), compiling a dataset on ministerial turnover and resignation calls.
- Data mining (OCR) and machine learning and then a semiparametric approach of competing risks and PS matching.

Theory and Empirical Expectations

Protection Policy and Dismissal Rule

To limit agency problems, the president could apply a combination of protection policy and a dismissal rule, borrowed from the rules modelled for parliamentary systems by Dewan and Myatt (2007, 2010; see also Berlinski et al., 2010).

The protection policy considers that ministers can be: (1) **tainted** by having been affected by a scandal or (2) have a record **clean of questioning**.

- If a tainted minister is again affected by a scandal, dismissal should be highly probable (limiting moral hazard).
- A minister with a clean record should be protected by the principal to encourage political activism (limiting agency loss).

Protection Policy Expectation

The first resignation call should serve as a warning for the minister to improve their performance rather than as an immediate threat to their position. This protection policy should have two outcomes: the retention of the questioned ministers or their reallocation to a different portfolio.

Thus, my first hypothesis is:

- **Protection Policy Hypothesis.** The first call for a minister's resignation raises the probability of reallocation, but not of individual terminations.

Dismissal Rule Expectation

If ministers whose resignation has already been called for are involved in a new scandal, the president could activate the dismissal rule.

The dismissal of a tainted minister could operate as a positive signal in the chain of delegation between voters and the president (Dewan and Dowding, 2005). In addition, dismissing tainted ministers could be a way to limit moral hazard by demonstrating that officeholders' actions are not insulated from negative consequences. Accordingly, the empirical expectation is:

- **Dismissal Rule Hypothesis.** The second and subsequent calls for a minister's resignation raise the probability of individual terminations, but not of reallocations.

Moderation Analysis

I focus on tainted ministers and test the reliability of the dismissal rule under two conditions: **nonpartisanship** and **seniority** of ministers.

First, the appointment of nonpartisan ministers indeed operates as a strategy for dealing with the agency problems, such as agency loss, information asymmetries and moral hazard, that can occur when the agent is a partisan minister (Altman, 2000; González-Bustamante, 2022; Martínez-Gallardo and Schleiter, 2015).

Second, by appointing and protecting more technically skilled cabinet members, presidents should achieve good public policy outcomes and send positive signals to the electorate (electoral accountatbility literature; see for example Besley and Case, 2003).

Empirical Strategy

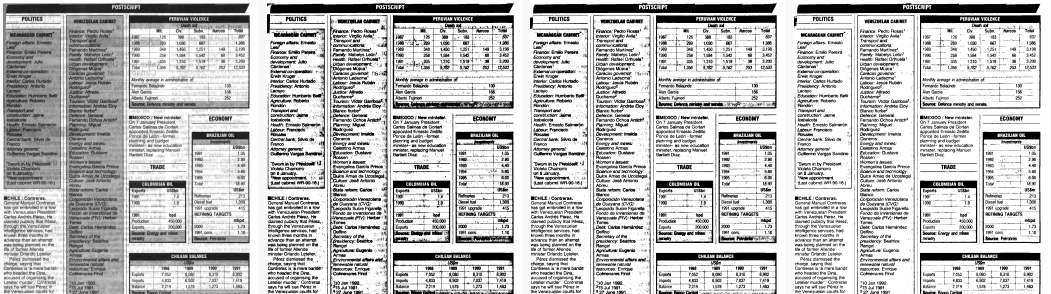
Cases and Data Gathering

I combined the Tesseract Optical Character Recognition (**OCR**; see Ooms, 2021) with different **semi-supervised machine learning models** to distinguish between resignation calls and other mentions of cabinet members in the press. This was applied to **46 years** of **Latin American Weekly Report (LAWR) archives**.

Detail of the Batches of LAWR Archives

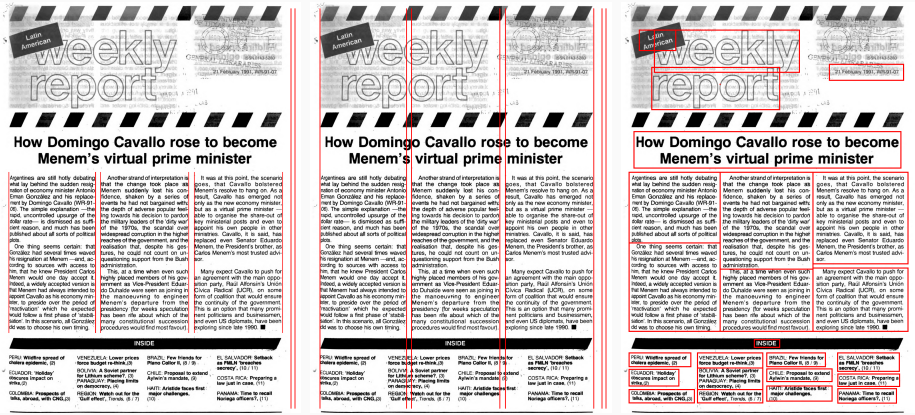
Batch	Archives	Years	Format	Volume
Batch-01	LAWR-1975 to LAWR-1979 Including 1997	5	CSV UTF-8	9.56 MB
Batch-02	LAWR-1980 to LAWR-1998 Excluding 1997 Including the first half of 2003	18.5	PNG images	38.5 GB
Batch-03	LAWR-1999 to LAWR-2021 Excluding the first half of 2003	22.5	PDF files	341 MB

Binarisation Techniques for Enhancing OCR



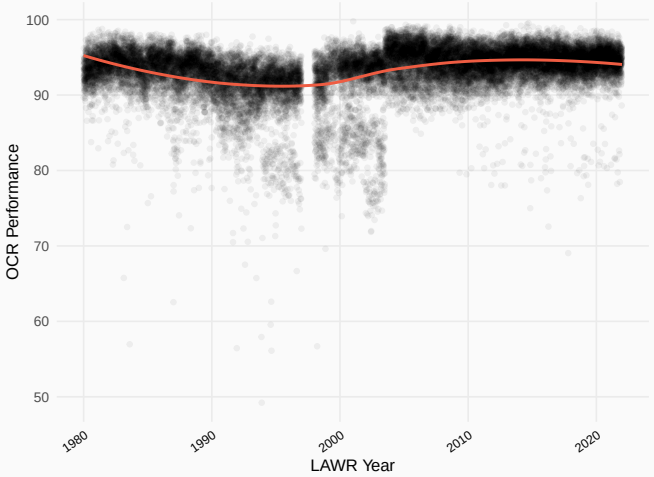
Note. The images correspond to a LAWR page before binarisation and after the application of the techniques of Sauvola and Pietikäinen (2000), Singh et al. (2011) and Su et al. (2010).

Tesseract Optical Character Recognition



Tab-Stop Lines — Column Layout — Segmented Blocks

Performance and Accuracy of the Tesseract OCR Algorithm



Accuracy is based on the proportion of text that was correctly identified.

This calculation was performed after extracting stop words and carried out with dictionaries trained with **Long Short-Term Memory (LSTM) models** used by Tesseract.

Semi-Supervised Approach

There are two significant **drawbacks** to training **supervised algorithms**:

1. This would have required a major amount of labelled text data **(time and resources)**
2. The accuracy of the algorithms is expected to decrease out-of-the-sample and over time **(batches of archives cover almost 50 years)**

Therefore, I preferred a **semi-supervised approach** that could maintain a **high accuracy level over time with a small amount of coded data** for training (Lee et al., 2021).

1. I drew a subsample of 1,000 LAWR press releases
2. A team of human coders classified all these observations (**resignation calls**)
3. Labelling was conducted to generate ground-truth labels to cross-validate
4. Semi-supervised models using knowledge-based seed words

Semi-Supervised Approach

Pre-processing

80% of the subsample for training

20% for validation

100 tokens per news item *

SMART stop words **

Customised list ***

Stemming algorithm

* until 500 for robustness checks

** Snowball for robustness checks

*** based on LAWR numbering

Models

Naïve Bayes (NB)

Support Vector Machine (SVM) *

Radial Basis Function (RBF) SVM **

Random Forest with 100 trees (RF100)

Random Forest with 500 trees (RF500)

Extreme Gradient Boosting (XGBoost)

* Kernel-based Linear

** Kernel-based Gaussian

Labelling Process for Benchmarking Supervised Models

Coder	Labels	$M(T)$ Label	$\sum$ Time	Consensus
1	966	39s	10h 25m	94.05
2	718	1m 15s	15h 0m	94.84
3	704	38s	7h 29m	94.82
4	318	1m 4s	5h 38m	95.47
5	263	17s	1h 14m	99.26
6	101	1m 17s	2h 9m	88.50

Note. Each observation was hand-coded by three different coders. 2.3% of observations were revised and relabelled during an iteration step. Seed words coincide with features extracted from the benchmarking supervised models trained based on this labelling process.

Goodness-of-Prediction of the Semi-Supervised Models

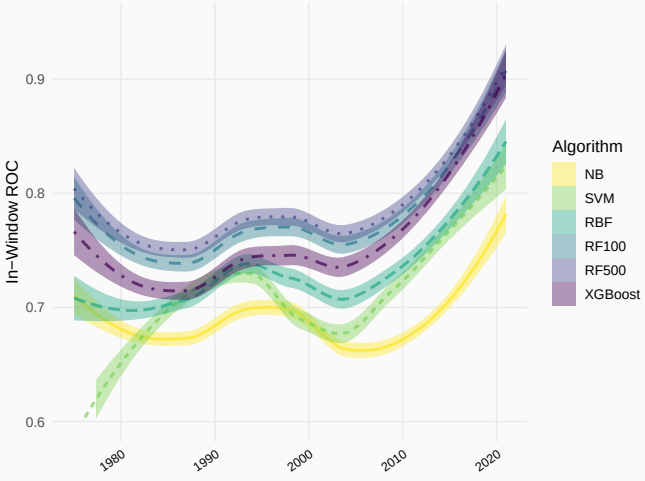
Model	Knowledge-Based Seed Words				
	Accuracy	Precision	Recall	F1-score	ROC
Naïve Bayes	0.713	0.734	0.948	0.828	0.638
Kernel Linear SVM	0.698	0.741	0.900	0.812	0.614
Kernel Gaussian RBF SVM	0.726	0.728	0.995	0.840	0.647
Random Forest (100 trees)	0.743	0.742	0.991	0.849	0.671
Random Forest (500 trees)	0.743	0.740	0.997	0.849	0.682
XGBoost	0.726	0.762	0.905	0.827	0.660

Note. The evaluation was performed with resampling methods using the testing set and 10-fold cross-validation.

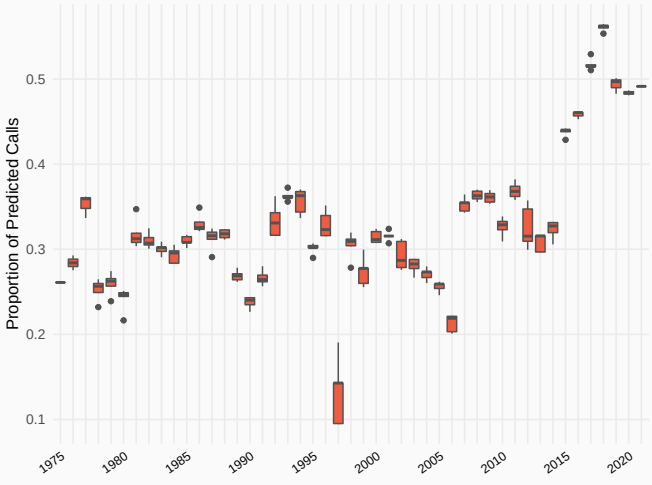
Goodness-of-Prediction of Semi-Supervised Models (1975-2021)

Following Greene et al. (2019), I trained semi-supervised models using a **five-year fixed rolling window** from 1975 to 2021 to train algorithms and predict resignation calls.

Random Forest classifiers with 500 trees 10-fold cross-validated.



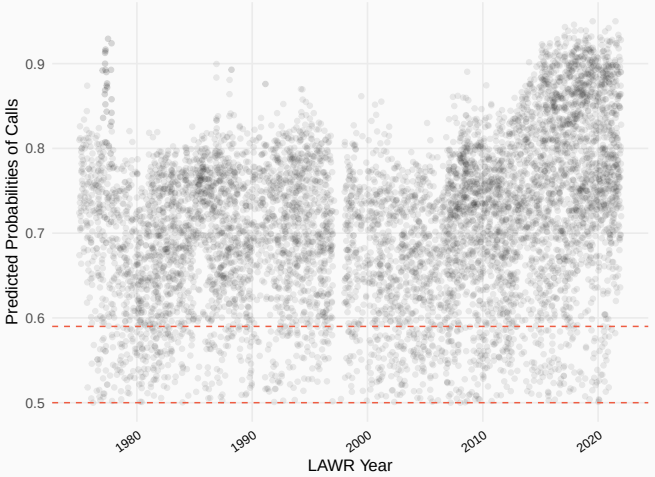
Predicted Ministerial Resignation Calls over Cabinet Press Mentions



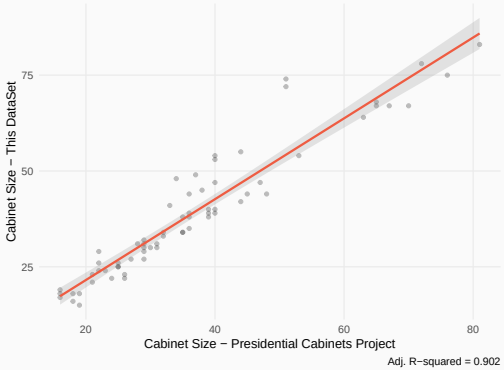
Proportion of predicted resignation calls over mentions of cabinet members in press reports, were carried out with an **ensemble semi-supervised RF500 algorithm**.

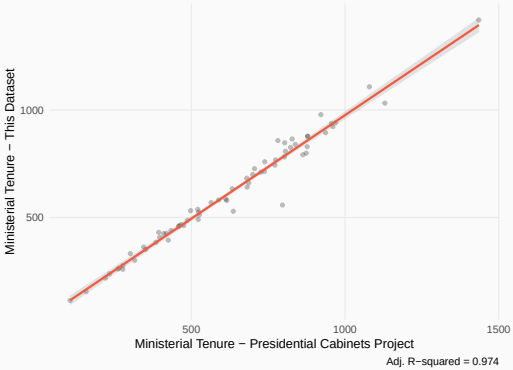
Augmented Artificial Intelligence (A2I)

I performed an augmented artificial intelligence process based on a **human review** of the predicted observations with low confidence values to improve the data's quality.



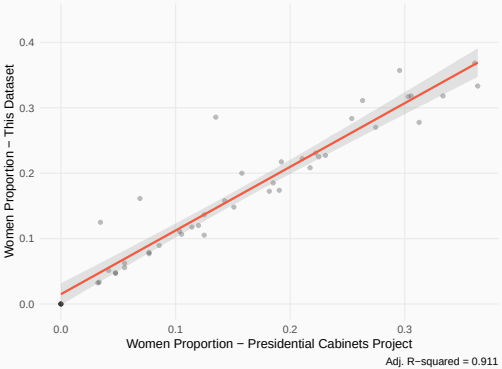
Convergent Validation against Similar Datasets

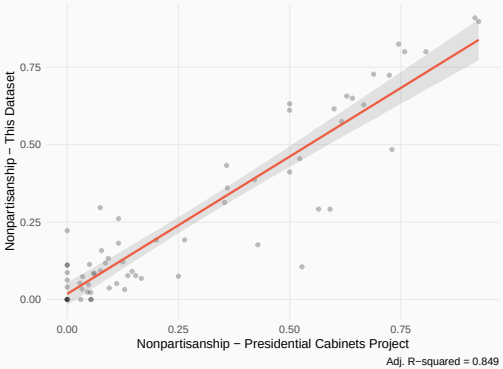




Note. Concurrent presidential terms were used for comparisons with the Presidential Cabinet Project (see Camerlo and Martínez-Gallardo, 2020).

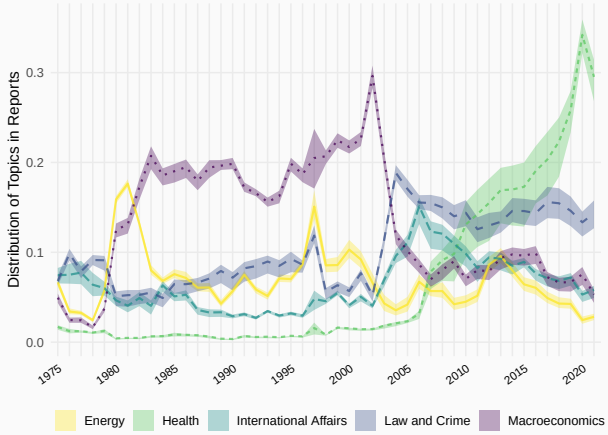
Convergent Validation against Similar Datasets





Note. Concurrent presidential terms were used for comparisons with the Presidential Cabinet Project (see Camerlo and Martínez-Gallardo, 2020).

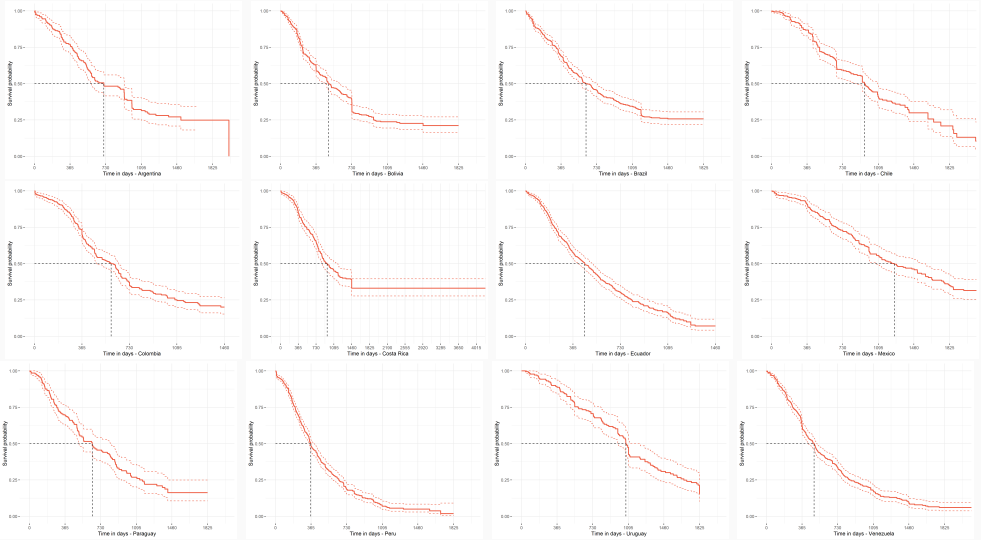
Content Validation: Major Agenda Topics in Latin America



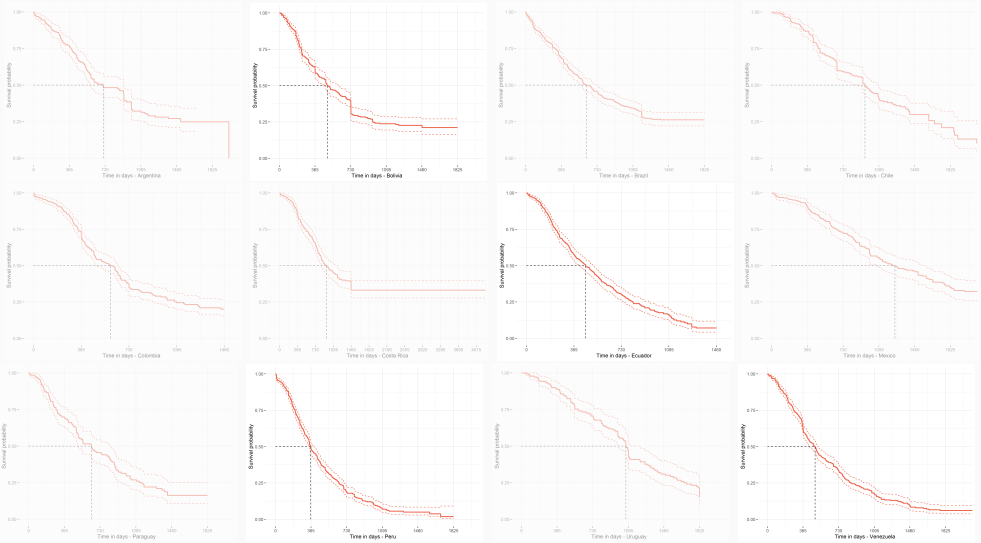
I used **Keyword Assisted Topic Modelling** (**keyATM**; see Eshima et al., 2022) and a dictionary based on Comparative Agendas Project (2022).

It is a semi-supervised Bayesian application that allows me to obtain ϑ (distribution of topics in the documents and CIs).

Kaplan-Meier Survival Estimations in 12 Presidential Cabinets

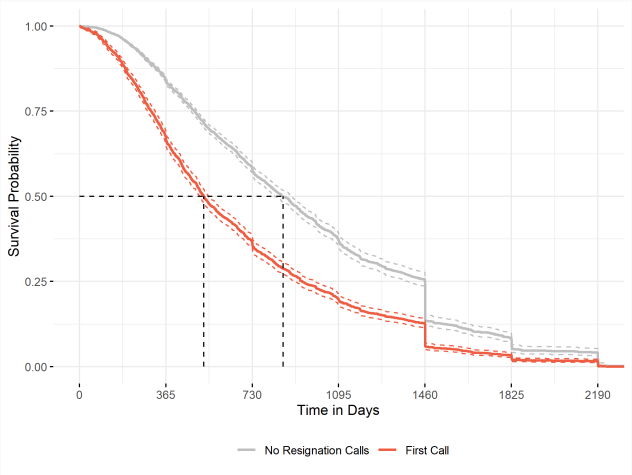


Kaplan-Meier Survival Estimations in 12 Presidential Cabinets



Kaplan-Meier Survival Estimation

$F(t) = 1 - S(t) = Pr(T \leq t)$ can be used to obtain the incidence of ministerial exits during T by considering $S(t)$ as a survival function, if raw exit is considered an event in the **absence of competing risks**.



Time-Dependent Data Encoding

The dataset was **time-dependent encoded** using monthly intervals with the (start, stop] model (Therneau et al., 2020). This allowed us to incorporate **time-varying covariates**.

We observe the time events T considering monthly intervals and k-*th* **competing risks** $Y_{k[i]}$ for **reallocations** ($k = 1$) and **individual ministerial terminations** ($k = 2$) by constructing the monthly intervals $Z(t) = I(t > Y_{k[i]})$.

I then used the closed interval of i-*th* observations (officeholders) as an endpoint in a function for multi-state survival variables obtaining an outcome dataset to control competing risks with **Fine-Gray weights** (Fine and Gray, 1999; Therneau et al., 2022).

I distinguish $D_{j[i]}$ as **1st resignation call** ($j = 1$) and the **2nd or subsequent ones** ($j = 2$). Consequently, I regress $D_{j[i]}$ on a vector of covariates and potential confounders $X_{m[i]}$ that considers quadratic presidential term patterns dummies and type of ministry, using probit models.

I also incorporated country FE and Fine-Gray weights w_1. This is our **naive propensity score estimation**:

$$D_{j[i]} = \varphi \left[\alpha \; + \sum_{m=1}^{M} \gamma_m w_{1[i]} X_{m[i]} + \zeta w_{1[i]} \, country_i + \varepsilon_i \right] \tag{1}$$

Propensity Score and NNM

Then, I expanded our PS estimation by incorporating **additional confounders** to $X_{m[i]}$ to **block the backdoor path**. I favour the incorporation of covariates that can affect the presidential decision (protect/dismiss), as well as calls:

- Presidential leverage (control of the relevant Houses)
- Opposition fragmentation*
- Government fragmentation*
- Government type (single-party vs. coalition)
- Presidential re-election
- Cumulative level of resignation calls**

$\star$ probability that two randomly selected deputies belong to different parties.

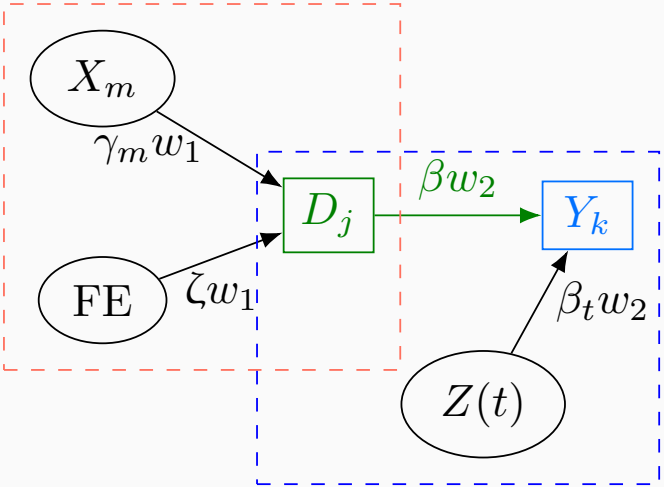
$\star\star$ to improve precision since it could be a competing exposure rather than a confounder.

After matching, the outcome analyses were carried out in the different matched samples with Fine-Gray subdistribution hazard models to estimate the **ATT for observational studies**, that is, the effect on ministers who received resignation calls.

The models regressed $Y_{k[i]}$, where $k = 1$ for **reallocations** and $k = 2$ for **dismissals**, on resignation calls indicators $D_{j[i]}$, using PS weights w_2 and $Z(t)$ intervals:

$$\lambda_k(t_i) = \lambda_{0[k]}(t_i)\ exp\ \left[\beta_t w_{2[i]} Z_i(t) + \beta w_{2[i]} D_{j[i]} + \varepsilon_i\right] \tag{2}$$

Nonparametric Graphical Representation



Main Results

Estimating the Protection Policy

	Reallocations		Ind. Terminations	
	Model I	Model II	Model III	Model IV
First resignation call	1.934*	3.304***	1.089***	1.628***
	(1.068)	(1.081)	(0.351)	(0.502)
Matching	NNM	NNM	NNM	NNM
Estimand	ATT	ATT	ATT	ATT
Competing risk	Ind. Term.	Ind. Term.	Realloc.	Realloc.
PS weights	Fine-Gray	Fine-Gray	Fine-Gray	Fine-Gray
Weights outcome	No	PS	No	PS
Nonpartisanship	No	PS	No	PS
Seniority	No	PS	No	PS
Control of Houses	No	PS	No	PS
Opp. fragmentation	No	PS	No	PS
Gov. fragmentation	No	PS	No	PS
Type of government	No	PS	No	PS
Re-election permitted	No	PS	No	PS
Cumulative gov. calls	No	PS	No	PS
Quadratic term patterns	PS	PS	PS	PS
Type of ministry	PS	PS	PS	PS
Country FE	PS	PS	PS	PS
Obs. clustering	PS/Yes	PS/Yes	PS/Yes	PS/Yes
Log-Rank	4.430**	4.069**	10.785***	12.213***
AIC	123.542	106.566	690.058	475.164
C-Index	0.676	0.690	0.614	0.653
N	10,143	9,900	8,884	6,580
Log Likelihood	−60.771	−52.283	−344.029	−236.582

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

I reject the **Protection Policy Hypothesis** as there is no evidence to suggest that the first call alone increases the probability of a minister being repositioned.

Although we reject this hypothesis, we found (novel empirical) evidence that reallocations and individual terminations are competing risks at the early stage of public questioning.

Estimating the Dismissal Rule

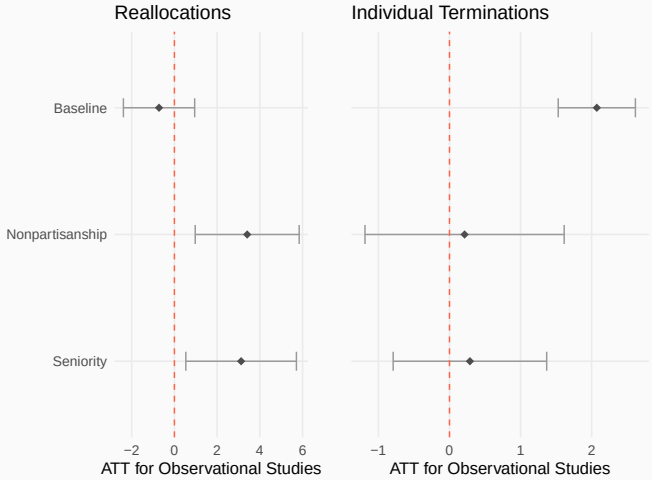
Model IV with the block of confounders has an estimated ATT of 2.073. This implies the risk of a minister's removal increases eightfold.

These results confirm the **Dismissal Rule Hypothesis**.

	Reallocations		Ind. Terminations	
	Model I	Model II	Model III	Model IV
Second or subsequent resignation calls	1.099 (0.744)	−0.729 (0.851)	1.329*** (0.283)	2.073*** (0.277)
Matching	NNM	NNM	NNM	NNM
Estimand	ATT	ATT	ATT	ATT
Competing risk	Ind. Term.	Ind. Term.	Realloc.	Realloc.
PS weights	Fine-Gray	Fine-Gray	Fine-Gray	Fine-Gray
Weights outcome	No	PS	No	PS
Nonpartisanship	No	PS	No	PS
Seniority	No	PS	No	PS
Control of Houses	No	PS	No	PS
Opp. fragmentation	No	PS	No	PS
Gov. fragmentation	No	PS	No	PS
Type of government	No	PS	No	PS
Re-election permitted	No	PS	No	PS
Cumulative gov. calls	No	PS	No	PS
Quadratic term patterns	PS	PS	PS	PS
Type of ministry	PS	PS	PS	PS
Country FE	PS	PS	PS	PS
Obs. clustering	PS/Yes	PS/Yes	PS/Yes	PS/Yes
Log-Rank	2.444*	4.386**	24.698***	35.444***
AIC	422.574	622.115	3,292.559	3,149.094
C-Index	0.566	0.595	0.573	0.595
N	30,422	31,250	21,511	21,905
Log Likelihood	−210.287	−310.058	−1,645.279	−1,573.547

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Moderation Analysis



The effect of resignation calls on the probability of reallocation is more substantial.

The estimated effects are not significant for individual termination in the case of nonpartisanship ($p = 0.766$) or seniority ($p = 0.602$), **confirming a protection strategy**.

Placebo Models

I repeated the primary models with resignation calls altered $\widetilde{D}_{j[i]}$ with the aim of applying a placebo test.

The **absence of a significant ATT** in most models tends to confirm the main results.

	Reallocations		Ind. Terminations	
	Model I	Model II	Model III	Model IV
Placebo	−0.129		0.678	
(press mentions)	(0.766)		(0.480)	
Placebo		−0.543		−0.627
(unrelated press mentions)		(0.523)		(0.385)
Matching	NNM	NNM	NNM	NNM
Estimand	ATT	ATT	ATT	ATT
Competing risk	Ind. Term.	Ind. Term.	Realloc.	Realloc.
PS weights	Fine-Gray	Fine-Gray	Fine-Gray	Fine-Gray
Weights outcome	PS	PS	PS	PS
Nonpartisanship	PS	PS	PS	PS
Seniority	PS	PS	PS	PS
Control of Houses	PS	PS	PS	PS
Opp. fragmentation	PS	PS	PS	PS
Gov. fragmentation	PS	PS	PS	PS
Type of government	PS	PS	PS	PS
Re-election permitted	PS	PS	PS	PS
Cumulative gov. calls	PS	PS	PS	PS
Quadratic term patterns	PS	PS	PS	PS
Type of ministry	PS	PS	PS	PS
Country FE	PS	PS	PS	PS
Obs. clustering	PS/Yes	PS/Yes	PS/Yes	PS/Yes
Log-Rank	0.271	7.501***	28.984***	78.112***
AIC	2,057.048	3,668.754	13,331.973	26,931.196
C-Index	0.523	0.530	0.534	0.553
N	108,493	180,699	81,736	134,976
Log Likelihood	−1,027.524	−1,833.377	−6,664.987	−13,464.600

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Robustness Checks

☑ The robustness checks using **additional confounders**, an **alternative matching algorithm** (caliper) and a **special standard error estimator** (Austin and Fine, 2019) show the same patterns as the main analyses.

☑ When incorporating macroeconomic indicators, presidential approval and controlling for low levels of freedom of speech as potential confounders, the results were not altered (**low residual confounding**).

☑ The exclusion of these **stochastic events** was made because I gave priority to the use of **resignation calls** as a noisy but empirically efficient indicator.

Takeaways and Future Work

Takeaways

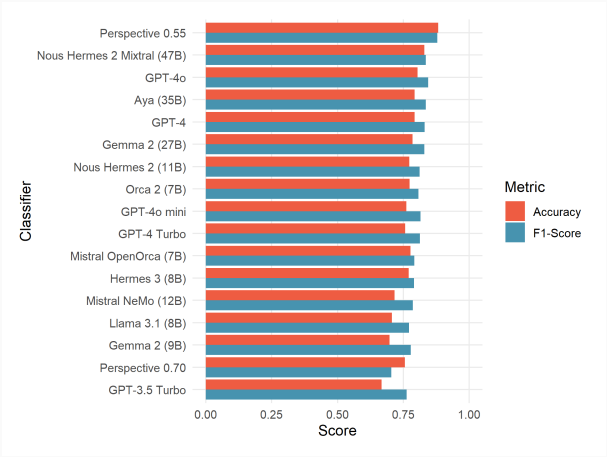
☑ I offer novel indicators such as **reallocations** and **resignation calls**. In addition, the use of semi-supervised ML models is also an innovation.

☑ The theoretical argument is simple: The president protects her ministers in the face of an initial questioning by reallocating them to a new ministry but removes them from the cabinet in the face of further resignation calls. **In a nutshell, it is like in the football**.

☑ Combining semiparametric competing risks models and propensity score contributes to tackling the **non-random selection problem** in observational data.

☑ The risk of ministerial terminations increases **eightfold** from the second call, but it is **moderated** in the case of nonpartisan and senior ministers.

Future Work: LLMs Classification



Perspective API using a laxer threshold, **GPT-4o**, and **Nous Hermes 2 Mixtral** outperform other LLM's zero-shot classification annotation for toxicity.

New avenues. More complex classifications (few-shot or chain-of-thought) and the use of synthetic data.

References

References I

Altman, D. (2000). The Politics of Coalition Formation and Survival in Multiparty Presidential Democracies: The Case of Uruguay, 1989-1999. *Party Politics*, 6(3):259–283. https://doi.org/10.1177/1354068800006003001.

Austin, P. C. and Fine, J. P. (2019). Propensity-score matching with competing risks in survival analysis. *Statistics in Medicine*, 38(5):751–777. https://doi.org/10.1002/sim.8008.

Berlinski, S., Dewan, T., and Dowding, K. (2010). The Impact of Individual and Collective Performance on Ministerial Tenure. *The Journal of Politics*, 72(2):559–571. https://doi.org/10.1017/s0022381609990843.

Besley, T. and Case, A. (2003). Political Institutions and Policy Choices: Evidence from the United States. *Journal of Economic Literature*, 41(1):7–73. https://doi.org/10.1257/002205103321544693.

Camerlo, M. and Martínez-Gallardo, C. (2020). The Presidential Cabinets Project. Research project, Universidade de Lisboa and University of North Carolina at Chapel Hill. Available at https://www.presidentialcabinets.org.

References II

Comparative Agendas Project (2022). Comparative Agendas Project (CAP). Comparing Policies Worldwide. Research project, Comparative Agendas Project. Available at https://www.comparativeagendas.net.

Dewan, T. and Dowding, K. (2005). The Corrective Effect of Ministerial Resignations on Government Popularity. *American Journal of Political Science*, 49(1):46–56. https://doi.org/10.1111/j.0092-5853.2005.00109.x.

Dewan, T. and Myatt, D. P. (2007). Scandal, Protection, and Recovery in the Cabinet. *American Political Science Review*, 101(1):63–77. https://doi.org/10.1017/S0003055407070025.

Dewan, T. and Myatt, D. P. (2010). The Declining Talent Pool of Government. *American Journal of Political Science*, 54(2):267–286. https://doi.org/10.1111/j.1540-5907.2010.00430.x.

Eshima, S., Sasaki, T., and Imai, K. (2022). Keyword Assisted Topic Models. Package, Harvard University and Massachusetts Institute of Technology. Package 'keyATM' version 0.4.1.

References III

Fine, J. P. and Gray, R. J. (1999). A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*, 94(446):496–509. https://doi.org/10.1080/01621459.1999.10474144.

González-Bustamante, B. (2022). Ministerial stability during presidential approval crises: The moderating effect of ministers' attributes on dismissals in Brazil and Chile. *The British Journal of Politics and International Relations*. OnlineFirst. https://doi.org/10.1177/13691481221124850.

Greene, K. T., Park, B., and Colaresi, M. (2019). Machine Learning Human Rights and Wrongs: How the Successes and Failures of Supervised Learning Algorithms Can Inform the Debate About Information Effects. *Political Analysis*, 27(2):223–230. https://doi.org/10.1017/pan.2018.11.

Lee, G. T., Kim, C. O., and Song, M. (2021). Semisupervised sentiment analysis method for online text reviews. *Journal of Information Science*, 47(3):387–403. https://doi.org/10.1177/0165551520910032.

Martínez-Gallardo, C. and Schleiter, P. (2015). Choosing Whom to Trust: Agency Risks and Cabinet Partisanship in Presidential Democracies. *Comparative Political Studies*, 48(2):231–264. https://doi.org/10.1177/0010414014544361.

References IV

Ooms, J. (2021). Open Source OCR Engine. Package, Berkeley Institute for Data Science. Package 'image.binarization' version 0.1.2.

Sauvola, J. and Pietikäinen, M. (2000). Adaptive document image binarization. *Pattern Recognition*, 33(2):225–236. https://doi.org/10.1016/S0031-3203(99)00055-2.

Singh, T. R., Roy, S., Singh, O. I., Sinam, T., and Singh, K. M. (2011). A New Local Adaptive Thresholding Technique in Binarization. *IJCSI International Journal of Computer Science Issues*, 8(6–2):271–277. https://doi.org/10.48550/arXiv.1201.5227.

Su, B., Lu, S., and Tan, C. L. (2010). Binarization of historical document images using the local maximum and minimum. In *DAS '10: Proceedings of the 9th IAPR International Workshop on Document Analysis Systems*, pages 159–166, Boston. ACM Press. https://doi.org/10.1145/1815330.1815351.

Therneau, T., Crowson, C., and Atkinson, E. (2020). Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model. Technical report, Mayo Clinic. Package 'survival' version 3.2-7.

References V

Therneau, T., Lumley, T., Atkinson, E., and Crowson, C. (2022). Kaplan-Meier and Aalen-Johansen (Multi-State) Curves, Cox Models, and Parametric Accelerated Failure Time Models. Technical report, Mayo Clinic. Package 'survival' version 3.4-0.

Thank you very much!

Do you have any questions?

b.a.gonzalez.bustamante@fgga.leidenuniv.nl

bgonzalezbustamante.com

Presentation compiled with LaTeX and some

Download the latest version from GitHub